



## Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at <http://about.jstor.org/participate-jstor/individuals/early-journal-content>.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

# THE MEASUREMENT OF WRITTEN COMPOSITION IN GRADES IV TO VIII

MATTHEW H. WILLING  
University School, Cleveland, Ohio

It was the writer's privilege two years ago to be associated with the survey staffs which investigated the work of the schools in Denver, Colorado, and in Grand Rapids, Michigan. His assignment in each case was the measurement of written composition in Grades IV to VIII inclusive. The purposes of the present article are to explain the method by which this assignment was met, and to indicate its possibilities for general use by teachers and supervisors.

At the outset of the writer's work these four principles were recognized as properly determinative of method in measuring written composition: (1) The accomplishment of pupils in written composition should be examined in their original work. (2) Spontaneity and interest in writing these compositions should be secured by assigning a subject of certain and wide suggestion to all the pupils. (3) The test should be administered in exactly the same way in all schools. (4) The papers should be graded by one person, or by persons of the same training in the use of a scale derived, preferably, from the test material itself. The actual procedure developed from these basic principles outlined itself under the following heads: (1) "Preparing the Pupils," (2) "Giving the Test," (3) "Deriving a Measuring Scale," and (4) "Grading the Papers."

## PREPARING THE PUPILS

Previous to beginning the work this notice was sent to the principals of the buildings concerned:

A special examiner will arrive at your building ..... the .....  
[time] [day] [date]  
at ..... for the purpose of giving a test in written composition. This test will be given to the pupils in Grades IV A, VA, VI A, VII A, and VIII A.

It will consist in having the pupils write original stories on a subject announced by the examiner. About thirty minutes will be consumed in each room, but the examiner will undertake to start the tests at five-minute intervals in the five rooms concerned, so that the whole time taken up in a building will not exceed one hour. The room teachers will be left in charge, under instructions, during the actual writing of the compositions. Preparation for this test involves only the making of a five-minute schedule for starting in the five rooms, and instructions to the teachers to see that pupils are provided with paper, pens, and ink, and that the papers are headed with name, age, school, and grade before the examiner arrives.

Avoid any discussion of the test with the pupils and all preliminary excitement of any kind whatsoever.

#### GIVING THE TEST

The writer, as examiner, upon reaching a building, received the five-minute schedule from the principal and at the proper time entered the first room. If the children were all ready, as was usually the case, he handed a paper to the teacher with the following written on it:

#### AN EXCITING EXPERIENCE

A Storm	In the Woods
An Accident	In the Mountains
An Errand at Night	On the Ice
A Wonderful Journey	On the Water
An Unexpected Meeting	A Runaway

All this he asked her to write on the board while he talked to the pupils.

His directions to the pupils ran in this wise—adapted, of course, to the maturity of the grade being tested:

I am here to have you write me a story. It is to be a story about some exciting experience that you have had, about something or other very interesting that has happened to you. If nothing of the sort has ever happened to you, then tell me of an exciting experience someone you know has had. You may even make up a story of this kind, if you have to; though I believe you will do better, on the whole, with a real one. I am going to give you about twenty minutes in which to write. During this time I shall be out of the room and you will not be interrupted till I return. You are to write on both sides of the paper, to do all the work yourselves, and to ask no questions at all after you begin. You may make any corrections you wish between the lines. There will be no time to rewrite your story.

Your teacher has written the general subject on the board together with some suggestions. You do not have to write on any of these topics unless you want to; they are merely to help out in case you cannot think of exciting experiences yourself. You may begin now as soon as you wish.

Opportunity was always given for asking questions, and the effort made to put the children as much at their ease as possible before the examiner left the room. Then the room teacher was left in charge under injunction to answer no questions and to permit no interruptions till the examiner returned.

This introduction to the actual writing usually took from three to five minutes, depending somewhat on the grade. A full twenty minutes was allowed for the writing, at the end of which time the examiner returned and directed the pupils as follows: "You are to have four or five minutes in which to finish your stories, make corrections, and count the number of words written. Write this number at the end of your story." At the conclusion of five minutes the examiner collected the papers and took them with him.

The whole procedure seemed to work very satisfactorily and was followed consistently in every grade. Possibly it could be improved by giving the pupils printed copies of the subject and the suggestions so as to avoid the slight distraction caused by the teacher's writing on the board during the remarks of the examiner. The difficulty might also be obviated by having the material placed on the board before the examiner arrives, but in a place concealed from the pupils.

The time allowed for the work proved sufficiently long as a rule. In the fourth and fifth grades it was ordinarily more than was necessary. Different classes, however, acted very differently in this respect. Some fourth grades found it hard to say all that was on their minds in the twenty-odd minutes, while some eighth grades exhausted themselves in fifteen minutes.

The suggesting of the topics seemed a very helpful device. Again and again the puzzled expressions on pupils' faces disappeared as their eyes ran down the list and they were reminded of this or that experience or escapade. On the other hand, the children did not by any means feel confined to the short list on the board. It served merely to save a great deal of time for many, and in this

way to invalidate the possible criticism that opportunity was not afforded for preliminary reflection and planning.

#### DERIVING A MEASURING SCALE

The method of deriving the measuring scale for these compositions was similar to that which has been employed in the derivation of various other school-subject scales now in use, particularly the Ayres scales for handwriting and spelling. It is a method based primarily on the assumption that the trait under consideration is distributed in accordance with the law of probability—in this case that merit in English composition among elementary-school children is distributed normally. This means, for example, that if a considerable number of the compositions secured as above and representing the different grades equally should be selected at random and their quality rated by some purely objective scale, one could predict from his knowledge of the probability curve the number of compositions rated at any point on this scale. Thus, if this objective scale discriminated only eight equally separated degrees of merit, the percentage of compositions of each degree, beginning with the lowest, would be about as shown in Table I.

TABLE I

	DEGREE							
	1	2	3	4	5	6	7	8
Percentage of compositions.....	4	9	16	21	21	16	9	4

We might then, if we so wished, take one composition from each of these quality groups and from them construct another scale. It is something of this sort that the writer did in the present instance.

This whole procedure can be carried out in a fairly valid way in the cases of certain kinds of material, such as handwriting and spelling, but requires considerable special adjustment in the case of written composition. Here there is no objective way in which to determine what the distribution of merit is. The writer's course was to select from the whole set a number of samples which

in his judgment did represent in themselves the assumed normal distribution of all. He accomplished this by a series of readings, groupings, and random selections too lengthy to detail here. Eventually he obtained a set of sixty-three samples which in merit ranged, according to his own standards of judgment, from the poorest to the best of the source material, and which, also in his judgment, distributed normally as explained above.

Specific care was taken in this selection that the samples should exhibit a high degree of correlation between their general merit and their formal correctness. By formal correctness here is meant correctness in spelling, punctuation, and grammar. This was ascertained through actual count of errors.

The sixty-three samples so chosen were then typewritten with all formal errors corrected, and submitted to twelve judges with instructions to rank them from poorest to best on purely rhetorical grounds. These rankings were then assembled and a composite ranking (ranking of average ranks) was derived, which determined each sample's exact place in the assumed normal distribution. Finally eight samples were selected to represent the extremes of this distribution and six equally separated points within. For those acquainted with statistical methods it may be said that the samples selected were such as approximately coincided in position with ordinates raised at sigma distances on the base of the error curve, seven such distances being regarded as including, for practical purposes, the whole distribution. The actual rankings of the samples so chosen were as follows: 1, 7, 14, 26, 40, 51, 59, and 63. To be statistically perfect these rankings should approach the center from the extremes by equal differences. The fact that they do not do so exactly is because the final selection of a sample for the scale took into consideration, not only its place in the distribution, but also its rank in the matter of formal correctness and the average variation of the judges in ranking it. Of two or three practically acceptable samples clustering about a point in the distribution, one was always chosen which showed the highest correlation between rhetorical and formal rankings and low variation in the estimates of the judges. The scale as finally made up accompanies this article.

## GRADING THE COMPOSITIONS

No composition scale yet published is objective either in its derivation or in its use. All are the product, more or less, of massed opinion, and in their application are at the mercy of the special intelligence and experience of those who use them. When we consider the degree of subjectivity which must enter into the comparison of samples of such a complex product as school composition with scale samples, it is not at all strange that hastily conducted experiments with such scales as this have shown little or no advantage in their use over unaided personal judgment. The value of composition scales can be ascertained validly only after the technique of their use is learned through intelligent practice. The writer's own experience is that the continued use of such aids in measuring compositional ability does very materially standardize his judgment. In his case repeated measurements of even small sets of papers, with sufficient time intervening to insure forgetfulness, have revealed practically no change in averages. Results of this sort are almost unknown in the case of uncontrolled personal judgment.

In the use of this particular composition scale, the writer has for himself developed the following technique: A composition is read carefully with attention both to rhetorical and to what we have been calling "formal" elements. As the reading progresses there is a conscious effort to place it on the scale, so that by the end of the reading its fate has frequently been decided. But in many, many cases the discrepancy between the story value and the form value is such as to dictate an adjustment or compromise of some sort. In these cases it is the writer's habit to assign the story value first and then to try to locate the formal worth. A compromise between the two is then assigned as the grade of the paper. Thus a composition grading very low as a story may by excellence in formal respects achieve as high a mark as 60; and, vice versa, a composition of high story value but low formal quality may be marked as low as 40. But no paper is marked above 70 which does not have both good story value and technical excellence, nor is a paper marked below 40 which does not lack both of these qualities. It is not possible to state exactly the relative emphasis placed on

story value and form value, but the effort is made, within the limits just mentioned, to keep the two approximately equal.

## STANDARDS

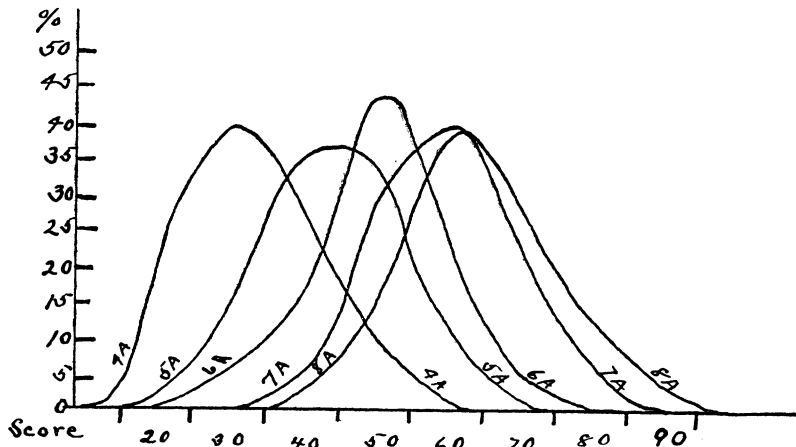
Extended discussions of the results of the composition surveys conducted by the writer in Denver and Grand Rapids may be found in the respective reports of the general investigations of those school systems. The averaged medians of the two cities as given in Table II may serve, however, as tentative standards for

TABLE II

COMPOSITION MEDIANS, AVERAGES OF DENVER AND GRAND RAPIDS MEDIANS

	GRADE				
	IV A	V A	VI A	VII A	VIII A
Quality.....	36.7	46.0	51.9	60.8	64.3
Quantity (words written) .....	103	137	158	182	197

those desiring to measure the work in their own schools. The medians in Table II represent an achievement about eight weeks before the end of the grade year.



Distribution curves of compositional merit, by grades. Denver, Colorado, December, 1915.

The accompanying diagram shows the nature and the inter-relation of the quality distributions by grades in the case of Denver.



The writer is inclined to see at least two advantages in his scale over others, if used under conditions exactly similar to those he has described. In the first place it is made up of the same material it attempts to measure. All the samples are narratives recounting, or attempting to recount, exciting experiences. Comparison of similar compositions with them is, practically speaking, easier than with samples less limited in range.

In the second place, attention has been paid in the make-up of the scale to the correlation of rhetorical and formal qualities in the samples. This makes it possible to give each factor more definite consideration in arriving at a final judgment of a composition's worth, an advantage that may easily be extended by an actual count of errors on the basis of some arbitrarily arranged key, such as those reported in the *Sixteenth Yearbook of the National Society for the Study of Education*, Part I.

In common with other composition scales it shares the disadvantages of subjective derivation, of unproved underlying statistical assumptions, of necessary training in the technique of its use, and of relative uselessness for individual diagnosis. The recognition of these defects, however, should not blind teachers and supervisors of English to the very considerable benefits to be obtained from the use of this and other similar scales for the measurement of general merit in school composition. It may well be indeed, even after the measurement of English composition shall have been placed on a thoroughly objective and analytical basis, that the scale for general merit will persist for the training of individual judgment and for ascertaining compositional situations in the gross.<sup>1</sup>

#### SCALE FOR MEASURING WRITTEN COMPOSITION IN GRADES IV TO VIII

THE VALUES 20, 30, 40, 50, 60, 70, 80, AND 90 GIVEN THE RESPECTIVE SAMPLES  
ARE ARBITRARY AND MERELY FOR PRACTICAL CONVENIENCE;  
20 MEANS 15 TO 24.9, 30 MEANS 25 TO 34.9, ETC.

#### 20

Deron the summer I got kicked and sprain my arm. And I was in bed  
of wheeks And it happing up to Washtion Park I was going to catch some fish.  
And I was so happy when I got the banged of I will nevery try that stunt againg.

Number of mistakes in spelling, punctuation, and syntax per hundred words, 30.

<sup>1</sup> See Monroe, De Voss, and Kelly, *Educational Tests and Measurements*, for comparative presentation of composition scales, including that described above.

## 30

The other day when I was rideing on our horse the engion was comeing and he got frightened so he through me down and I broke my hand.

And the next thing I done was I went to the doctor and he put some bandage on it and told me to come the next day so I came the next day and he toke the bandage off and he look at it and then it was better.

---

Number of mistakes in spelling, punctuation, and syntax per hundred words, 23.

## 40

My antie had her barn trown down last week and had all her chickens killed from the storm. Whitch happened at twelve oclock at night. She had 30 chickens and one horse the horse was saved he ran over to our house and claped on the dor whit his feet. When we saw him my father took him in the barn where he slepped the night with our horse. When our antie told us about the accident we were very sorry the next night all my anties things were frozen. The storm blew terrible the next morning and I could not go to school so I had to stay home the whole week.

---

Number of mistakes in spelling, punctuation, and syntax per hundred words, 17.

## 50

One time mother and father were going to take sister and I for a long ride thanksgiving. We had to go 60 miles to get there. When sister and I herd about it we were very glad. It was a very cold trip. We four all went in a one seated automobile. Dady drove and mother held me and sister sat on the top the top was down. Mother could not hold sister for she was two heavy. When we got there they had a hot fire ready for us and a goose dinner. We were there over night. In the morning it was hot out. This was on a farm. Sister and I got to go horse-back riding. It was lots of funs. They had children. The children were very nice. Our trip home was very cold. When we got home it had snod.

---

Number of mistakes in spelling, punctuation, and syntax per hundred words, 14.

## 60

One time when mother, some girl friends and myself were staying up in the mountains. An awful storm came up. At the we were way up the mountain. The lightning flashed and the thunder roared. We were very frightened for the cabin we were staying at was at the foot of the mountain. We didn't have our coats with us for it was very warm when we started. There were a few pine trees near us so we ran under them. They didn't do much good for the rain came down in torrents. The rain came down so hard that it uprooted one of the trees. Finely it began to slack a little, So we thought we would try and go back. About half way down the mountain was a little hut. We started and when got about half way down it began to rain all the harder. We didn't know what to do for this time there wasn't any trees to get under. We decided to go on for the nearest shelter was the hut. Finely we got there cold and wet to the skin.

---

Number of mistakes in spelling, punctuation, and syntax per hundred words, 11.

## 70

When I was in Michigan I had an exciting thing happen or rather saw it, it was when the big steamship plying between Chicago and Muskegon was sunk about 7 o'clock in the evening. It caught on fire with a load of cattle and products from the market on board, one of the lifeboats carrying some of the people who were on board landed at our pier. The "Whaleback" steamer which goes between Chicago and Muskegon was two hours later in coming than the freighter and was stopped to clear up the wreckage. all of the cattle and products and an immense cargo of coal were lost, but there were only two people lost. the ship tried hard to get to port with her cargoe but, could not reach it. The next morning we found planks, and parts of the wreck on the beach. Our cottage was at the top of a cliff and it was just one hundred feet to the lake from our cottage, we had a beautiful view, and the sight of the fire on the horizon was a beautiful sight (though it was pitiful).

---

Number of mistakes in spelling, punctuation, and syntax per hundred words, 8.

## 80

Near our ranch in Fort Logan there was a chicken ranch. One day my sister and I went up to the chicken ranch on our horses. Coming back there was a road leading from our house to the main road and along this road were half rotted stumps. On every one of these stumps what do you think we saw. We saw snakes! snakes! snakes! I suppose these snakes were shedding their skins, they were of every color, shape, and size. But when sister and I saw these snakes we whipped our horses into a gallop and away we went just as hard as we could go. When we got to the house we went in and mamma couldn't get us out of the house that day. I was so scared that I believe I dreamed about snakes for a month.

---

Number of mistakes in spelling, punctuation, and syntax per hundred words, 5.

## 90

The most exciting experience of my life happened when I was but five years of age. I was riding my tricycle on the top of our high terrace. Beside the curbing below, stood a vegetable wagon and a horse. Suddenly I got too near the top of the terrace. The front wheel of my tricycle slipped over and down I went, licety-split, under the horse standing by the curbing. I had quite a high tricycle and the handlebars scraped the horse's stomach, making him kick and plung in a very alarming manner. I was directly under him during this, but finally rolled over out of his way and scrambled up. I looked at my hands! Most of the first finger and part of the thumb of my left hand were missing. The horse had stepped on them. I had endured no sensation of pain before this, but now my mangled hand began to hurt terribly. I was hurried to the hospital and operated on, and now you would hardly notice one of my fingers is missing. I certainly have good cause to congratulate myself on my good fortune in escaping with as little injury to myself as I did, for I might have been terribly mangled in my head or body.

---

Number of mistakes in spelling, punctuation, and syntax per hundred words, 0.